FIG. 1

a)

**Number of predicted ligands**

149

35

10    6    5    7    4    3    0    2    8    1    1

1    2    3    4    5    6    7    8    9    10    11 to 20    20-100    >100

**Number of isolates containing exact predicted ligand**

b)

**Number of predicted ligands**

420

190

51

30    26    10    18    2    4    0    25    7    7

1    2    3    4    5    6    7    8    9    10    11 to 20    20-100    >100

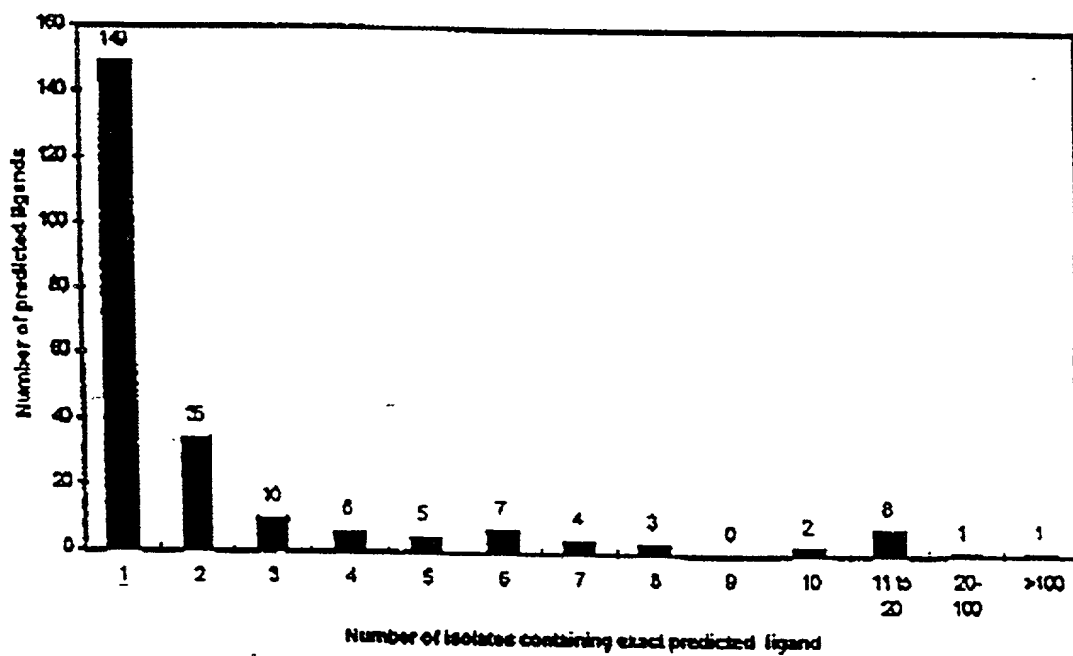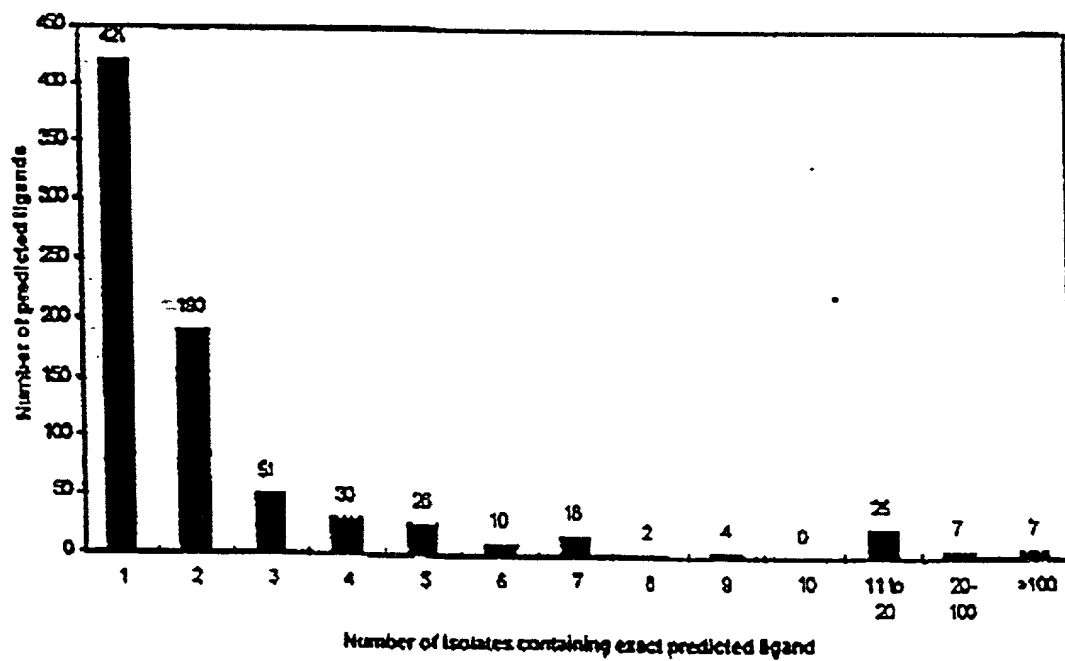**Number of isolates containing exact predicted ligand**

FIG. 2

| Sequence | A2 EBP | B27 EBP | A2 Fold Increase (less than 1.3 not reported) | B27 Fold Increase (less than 1.3 not reported) | Protein | Number of Isolates with Exact AA sequence | Approximate Position in LAI | Clade A | Clade B | Clade C | Clade D | Clade E | Other | SEQ ID NO. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KLTPLCVTLN | 55.68% | 0.00% | 1.33 | | Env | 159 | gp120 - 120 | X | X | X | X | X | X | 1 |
| AEWDRVHPV | 66.42% | 0.00% | 1.35 | | Gag | 36 | gag - 215 | X | X | X | X | X | X | 2 |
| SLFNTVATL | 62.00% | 0.00% | | | Gag | 18 | gag - 100 | X | X | X | X | | X | 3 |
| ELHPDKWTV | 57.03% | 0.00% | | | RT | 17 | RT - 354 | | X | | | | | 4 |
| GMDDPEREVL | 72.52% | 0.00% | 2.7 | | Nef | 17 | nef - 170 | | X | | | | | 5 |
| GMDDPEKEVL | 87.51% | 0.01% | 1.33 | | Nef | 16 | nef - 170 | | X | | | | | 6 |
| HLWRWGTMLL | 76.69% | 0.00% | | | Env | 10 | gp120 - 30 | | X | | | | | 7 |
| LLLTRDGGVN | 55.68% | 0.00% | | | Env | >10 | gp120 - 452 | X | | | X | | X | 8 |
| HLWKWSTMLL | 90.92% | 0.00% | 1.63 | | Env | >10 | gp120 - 20 | | | | X | | | 9 |
| ILKEPVHGV | 97.47% | 0.00% | 1.54 | | RT | >10 | RT - 480 | | X | | | | | 10 |
| KRWILGLNK | 0.00% | 14.22% | | 3.61 | Gag | 79 | gag - 263 | | X | | | | | 11 |
| CRIKQIIN | 0.00% | 99.08% | | | Env | 185 | gp120 - 420 | X | X | X | X | | X | 12 |
| CRIKQIINMW | 0.00% | 99.52% | 1.45 | 1.74 | Env | 150 | gp120 - 420 | X | X | X | X | | X | 13 |
| VSFEPIPIHF | 0.20% | 55.61% | | | Env | 109 | gp120 - 215 | | X | X | X | | X | 14 |
| RCSSNITGL | 0.01% | 62.11% | | | Env | 101 | gp120 - 446 | | X | | X | | X | 15 |
| VSFEPIPIHY | 0.00% | 98.22% | | | Env | 101 | gp120 - 215 | X | X | | X | | X | 16 |
| CRIKQIVNM | 0.00% | 91.33% | | | Env | 75 | gp120 - 420 | X | | X | X | | | 17 |
| IRSENITNN | 0.00% | 82.77% | | | Env | 42 | gp120 - 275 | X | | | | X | | 18 |
| IRIFMIV | 0.05% | 89.06% | | | Env | 19 | gp41 - 175 | X | | X | X | | | 19 |
| ISFDPIPIHY | 0.01% | 67.49% | | | Env | 15 | gp120 - 215 | | | | | X | | 20 |
| YRTGDIIG | 0.00% | 56.14% | | | Env | 15 | gp120 - 330 | | | | | X | | 21 |
| IRIGPGQTFY | 0.07% | 75.36% | | | Env | 13 | | | X | X | | | | 22 |
| GCSGKIIC | 0.00% | 61.09% | | | Env | 12 | gp41 - 90 | X | | | | X | | 23 |
| RRRAPQDS | 0.00% | 67.49% | | | Tat | 12 | | | X | | | | | 24 |
| IRSENITDN | 0.00% | 59.28% | | | Env | 11 | gp120 - 275 | | | | X | X | | 25 |
| CRIKQFIN | 0.00% | 76.92% | | 1.53 | Env | <10 | gp120 - 420 | | X | | | | | 26 |
| KRISIGPGR | 0.00% | 56.93% | | 1.78 | Env | <10 | gp120 - 320 | | X | | | | | 27 |
| GCQQIIEQL | 0.10% | 78.95% | | | Env | <10 | | X | | | | | | 28 |
| GRRGWEILKY | 0.01% | 59.80% | | 3.27 | Env | <10 | gp41 - 270 | | X | | | | | 29 |

FIG. 3

Project Outline

HLA allele selection

Indian HIV-1 sequences

EpiMatrix evaluation and selection

Binding assays (T2 stabilization)

CTL assays

Include CTL epitope in regional HIV-1 vaccine
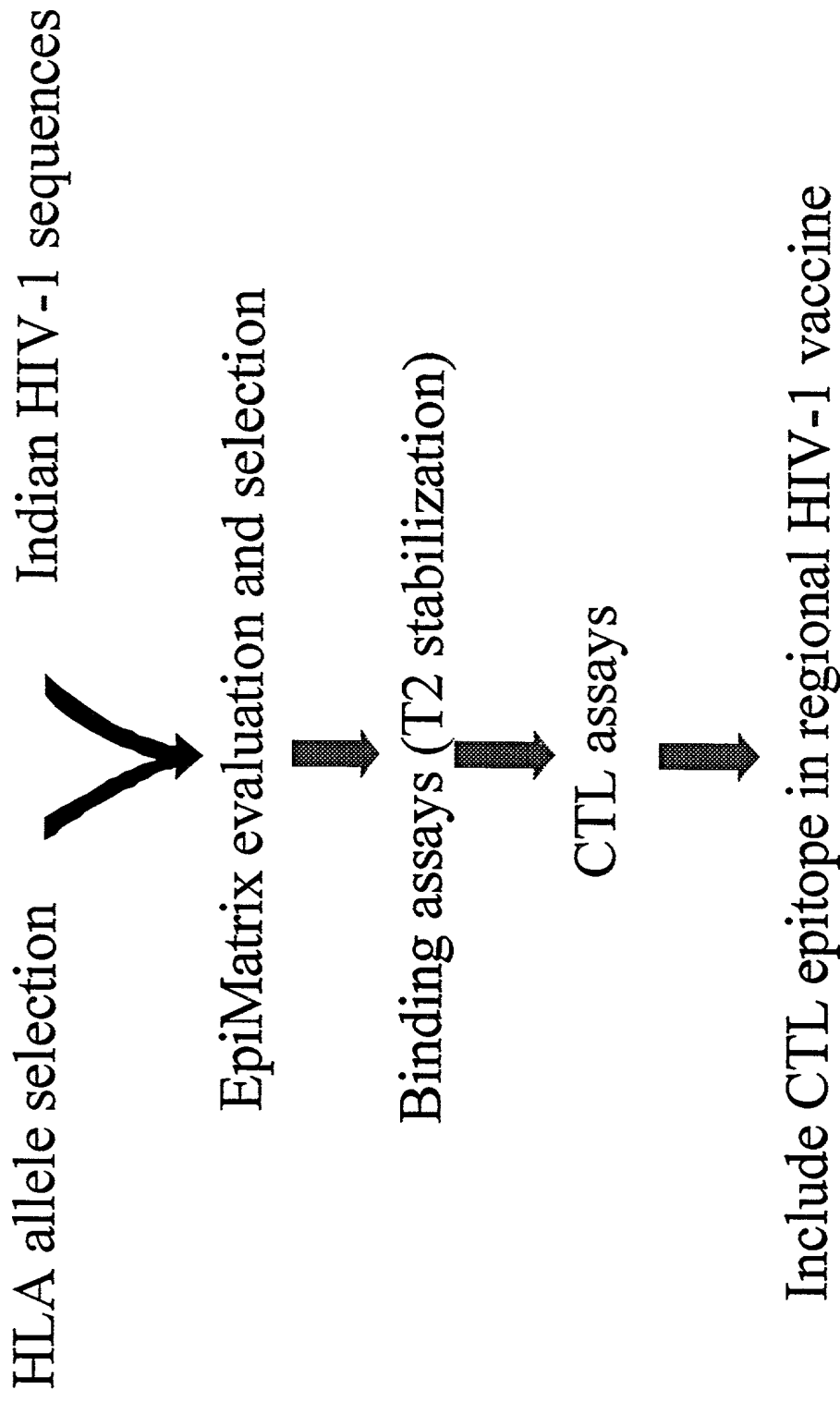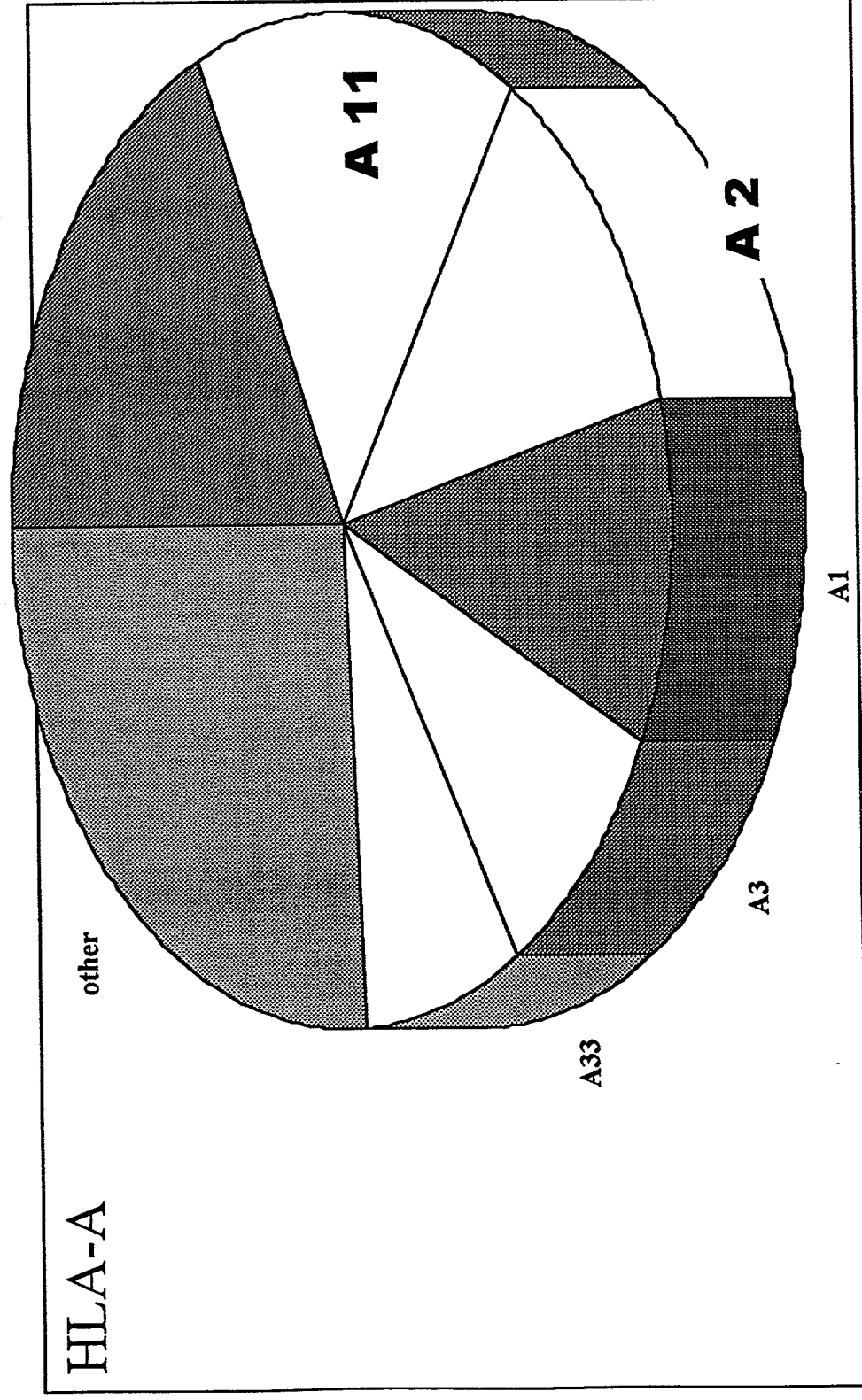
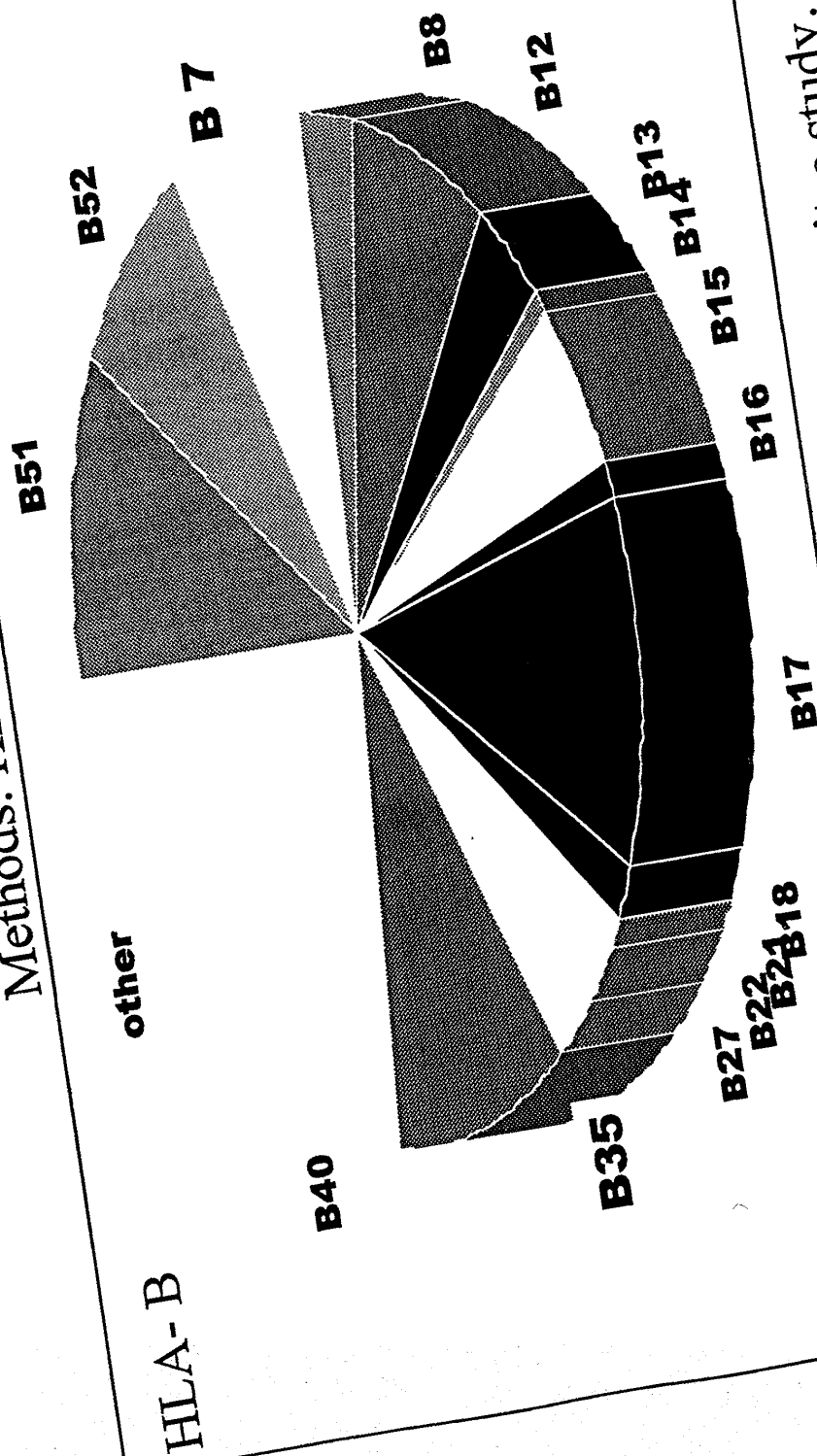## Methods: HLA allele selection



**HLA-A**

other

A33

A 11

A3

A1

A 2

HLA-A2 and -A11, -B7 and -B35 were selected for this pilot project due their prevalence of the HLA alleles in the Indian population….

FIG. 5

Methods: HLA allele selection



HLA- B

other

B51

B52

B 7

B8

B12

B13

B14

B15

B16

B17

B18

B21

B22

B27

B35

B40

... and availability of cell lines for *in vitro* study.

# FIG. 6

## EpiMatrix Predictions and Binding Results: B 7
### 6 out of 7, and control peptide

| B7 | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| peptide # | peptide | seq. Used | gene | strain | start-stop | % conserved | CTL | predicted EBP | avg MFI (200ug/ml) | avg fold incr. (20ug/ml) |
| 1 | RPNNNTRKSI | RPNNNTRKSI | ENV | DiD757 | 183-192 | 75 | Y | 8% | 335.6 | 2.4 |
| 3 | NPYNTPIFAL | NPYNTPIFAL | POL | SoInd5 | 61-70 | 60 | | 20% | 281.9 | 2.0 |
| 4 | RAIEAQQHLL | RAIEAQQHLL | ENV | DiD747 | 481-490 | 60 | | 17% | 181.5 | 1.3 |
| 5 | TCKSNITGLL | TCKSNITGLL | ENV | DiD760 | 375-384 | 59 | | 18% | 160.5 | 1.2 |
| | KPVVSTQLL | KPVVSTQLL | ENV | DiD747 | 182-191 | 71 | | 46% | 248.5 | 1.8 |
| | KPCVKLTPL | KPCVKLTPLC | ENV | DiD747 | 51-60 | 100 | | 27% | 373.8 | 2.7 |
| | GPKVKQWPL | GPKVKQWPLT | POL | SoInd4 | 25-34 | 100 | | 27% | 314.7 | 2.3 |
| | YPGIKVRQL | YPGIKVRQLC | POL | SoInd4 | 278-287 | 100 | | 26% | 378.4 | 2.7 |

# FIG. 7

## EpiMatrix Predictions and Binding Results: B 35
## 7 out of 7 … and control peptide

| B37 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| peptide # | peptide | seq. Used | gene | strain | start-stop | % conserved | CTL | predicted EBP | avg MFI (200ug/ml) | avg fold incr. (20ug/ml) |
| 2 | TVLDVGDAYF | TVLDVGDAYF | POL | SoInd4 | 114-123 | 100 | Y | 4% | 47.9 | 1.6 |
|  | EPPFLWMGY | EPPFLWMGYE | POL | SoInd4 | 231-239 | 100 |  | 9% | 48.7 | 1.6 |
| 6 | VPVKLKPGM | VPVKLKPGMD | POL | SoInd4 | 15-24 | 100 |  | 9% | 53.3 | 1.7 |
| 7 | CPKVTFDPI | CPKVTFDPIP | ENV | DID760 | 144-153 | 53 |  | 7% | 35.0 | 1.2 |
| 8 | KPVVSTQLL | KPVVSTQLL | ENV | DID747 | 182-191 | 71 |  | 9% | 40.5 | 1.4 |
|  | KPCVKLTPL | KPCVKLTPLC | ENV | DID747 | 51-60 | 100 |  | 11% | 52.1 | 1.7 |
|  | GPKVKQWPL | GPKVKQWPLT | POL | SoInd4 | 25-34 | 100 |  | 11% | 41.2 | 1.4 |
|  | YPGIKVRQL | YPGIKVRQLC | POL | SoInd4 | 278-287 | 100 |  | 7% | 40.7 | 1.3 |

# FIG. 8

## EpiMatrix Predictions and Binding Results: A 2
### 3 out of 7 … and control peptide

**A2**

| peptide # | peptide | seq. Used | gene | strain | start-stop | % conserved | CTL | predicted EBP | avg MFI (200ug/ml) | avg fold incr. (20ug/ml) |
|---|---|---|---|---|---|---|---|---|---|---|
| 13 | ILKEPVHGV | ILKEPVHGVY | POL | Solnd4 | 316-325 | 80 | Y | 96% | 1604.2 | 1.6 |
| 14 | QLPEKDSWTV | QLPEKDSWTV | POL | Solnd4 | 252-261 | 100 | | 87% | 1368.1 | 1.4 |
| 15 | NLWTVYYGV | NLWTVYYGV | ENV | GrD1024 | 32-41 | 67 | | 84% | 1716.9 | 1.8 |
| 16 | QMHEDVISL | QMHEDVISLW | ENV | DiD747 | 37-46 | 91 | | 78% | 1413.1 | 1.4 |
| 17 | KIEELREHLL | KIEELREHLL | POL | Solnd5 | 208-217 | 60 | | 79% | 889.9 | 0.9 |
| 18 | DMVNQMHEDV | DMVNQMHEDV | ENV | DiD747 | 33-42 | 64 | | 77% | 731.1 | 0.4 |
| 19 | GLKKKSVTV | GLKKKKSVTV | POL | Solnd4 | 106-115 | 100 | | 76% | 1088.4 | 1.1 |
| 20 | ELHPDKWTV | ELHPDKWTVQ | POL | Solnd4 | 240-249 | 80 | | 72% | 1048.1 | 1.0 |

# FIG. 9

## EpiMatrix Predictions and Binding Results: A11
### 4 out of 7 .... and control peptide

**A11**

| peptide # | peptide | seq. Used | gene | strain | start-stop | % conserved | CTL | predicted EBP | avg MFI (200ug/ml) | avg fold incr. (20ug/ml) |
|---|---|---|---|---|---|---|---|---|---|---|
| 21 | IYQEPFKNLK | IYQEPFKNLK | POL | SoInd4 | 348-357 | 100 | Y | 7% | 677.5 | 3.1 |
| 22 | VTFDPIPIHY | VTFDPIPIHY | ENV | DiD760 | 147-156 | 53 | | 22% | 190.0 | 0.9 |
| 23 | TVQCTHGIK | TVQCTHGIKP | ENV | DiD747 | 174-183 | 59 | | 44% | 733.4 | 3.3 |
| 24 | NTPIFALKKK | NTPIFALKKK | POL | SoInd5 | 64-73 | 60 | | 44% | 187.8 | 0.9 |
| 25 | LVDFRELNK | LVDFRELNKR | POL | SoInd4 | 81-90 | 100 | | 47% | 755.2 | 3.4 |
| 26 | PGMDGPKVK | PGMDGPKVKQ | POL | SoInd4 | 21-30 | 100 | | 52% | 193.8 | 0.7 |
| 27 | GIPHPAGLKK | GIPHPAGLKK | POL | SoInd4 | 100-109 | 100 | | 62% | 309.6 | 1.4 |
| 28 | FTTPDKKHQK | FTTPDKKHQK | POL | SoInd4 | 221-330 | 100 | | 63% | 920.6 | 4.1 |

# FIG. 10

## Methods: T2 Binding Assay

Allele matched peptides stabilize MHC molecules on the surface of TAP deficient cells. The stabilized MHC-peptide complex is detected using Ab to the MHC and fluorescence labeled secondary Ab.
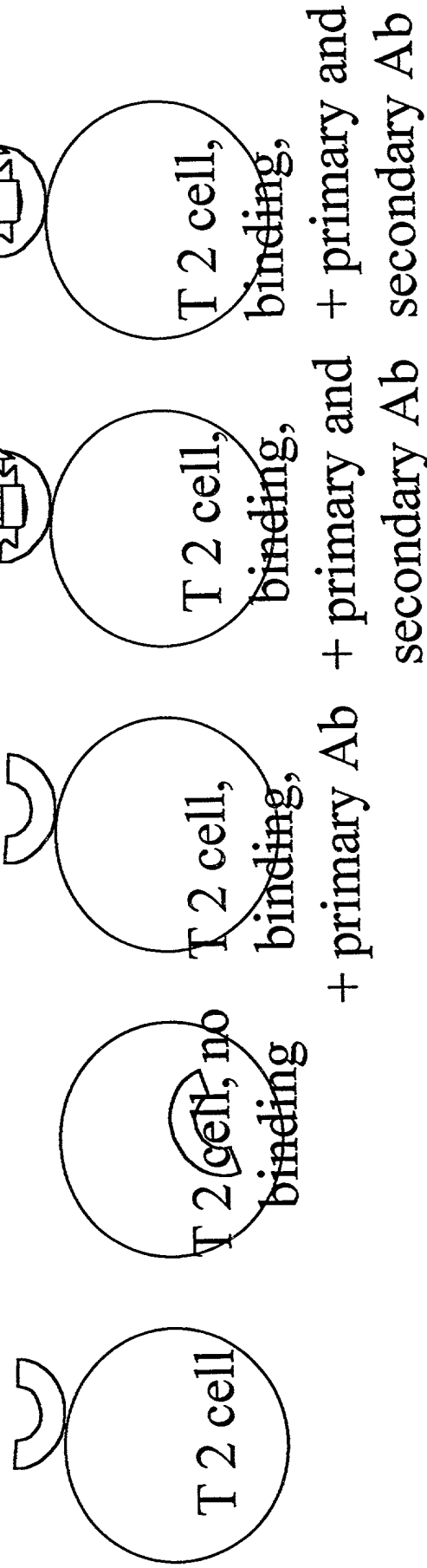
FIG. 11
Clustering of putative MHC ligands in *env*



Putative Ligand Location in LANL "most likely" sequence